

Data Visualization with R (ggplot2) vs. Matlab

	R (mainly ggplot2)	Matlab
Basics		
Import Dataset	<code>df = read.csv('StudentsPerformance.csv')</code>	<code>T = readtable('StudentsPerformance.csv')</code>
Column Names	<pre>>> colnames(df) [1] "gender" [2] "race.ethnicity" [3] "parental.level.of.education" [4] "lunch" [5] "test.preparation.course" [6] "math.score" [7] "reading.score" [8] "writing.score"</pre>	<pre>>> T.Properties.VariableNames {'gender'} {'race_ethnicity'} {'parentalLevelOfEducation'} {'lunch'} {'testPreparationCourse'} {'mathScore'} {'readingScore'} {'writingScore'}</pre>
Extract Data	Column: <code>df\$lunch; df['lunch']</code> Dataframe: <code>df[1:3, c('lunch','math.score')]</code> Filtering rows: <code>df[df\$math.score < 50,]</code>	Column: <code>T.lunch;</code> Table: <code>T(1:3, ['lunch','mathScore'])</code> Filtering rows: <code>T(T.mathScore<50,:)</code>
Convert Type	<code>as.factor(df\$gender)</code>	<code>categorical(T.gender)</code>
One Variable		
Histogram	<code>ggplot(df,aes(math.score))+ geom_histogram(bins=30, alpha=0.6, fill='cyan', color='blue')</code>	<code>histogram(df.mathScore,30,'FaceAlpha',0.6, ... 'FaceColor','cyan','EdgeColor','blue')</code>
Bar plot (counts)	<code>ggplot(df,aes(gender))+ geom_bar()</code>	<code>histogram(T.gender)</code>
Density Plot	<code>ggplot(df,aes(math.score))+ geom_density()</code>	<code>histfit(T.mathScore)</code>
QQ-plot	<code>ggplot(df)+ geom_qq(aes(sample=math.score))</code>	<code>qqplot(T.mathScore)</code>
Boxplot	<code>ggplot(df,aes(math.score))+ geom_boxplot() ggplot(df,aes(math.score, race.ethnicity))+ geom_boxplot()</code> (By race.ethnicity)	<code>boxplot(T.mathScore)</code> By race_ethnicity: <code>boxplot(T.mathScore, T.race_ethnicity)</code>
Two Variables		
Bar plot	<code>ggplot(df[1:3,], aes(1:3, math.score))+ geom_col()</code>	<code>bar(1:3, T.mathScore{1:3})</code>
Line plot	<code>ggplot(df1, aes(Date,Open))+ geom_line()</code>	<code>plot(T1.Date, T1.Open)</code>
Scatterplot	<code>ggplot(df,aes(math.score,reading.score))+ geom_point(color='blue', shape='.'))</code>	<code>scatter(T.mathScore,T.readingScore, 'blue', '.')</code>
Heatmap	<code>ggplot(df,aes(math.score,reading.score))+ geom_bin2d()</code>	<code>heatmap(T,'mathScore', 'readingScore')</code>
Other		
Parallel Coord	<code>library(GGally) ggparcoord(df, columns=6:8,alpha=0.4)</code>	<code>parallelplot(T(:,6:8))</code>
Map	<code>library(ggmap) ny_map <- get_stamenmap(bbox = c(left = -74.2591, bottom = 40.4774, right = -73.7002, top = 40.9162), zoom = 10, maptype = "toner-lite") ggmap(ny_map) + geom_segment(x=-73.9857,y=40.7484, xend=-74.1111, yend=40.7588,color='blue')</code>	<code>geoplot([40.7484,40.7588], ... [-73.9857,-74.1111], 'g-*') geolimits([40.4774, 40.9162], ... [-74.2591,-73.7002])</code>
Labels & Title	<code>ggplot(df,aes(math,,reading))+geom_point() + labs(x='math',y='reading',title='scores distribution')</code>	<code>xlabel('math') ylabel('reading') title('scores')</code>